AD_____

AWARD NUMBER DAMD17-97-1-7293

TITLE: Use of Genetic Suppressor Elements to Identify Potential Breast Tumor Suppressor Genes

PRINCIPAL INVESTIGATOR: Marija Zeremski

CONTRACTING ORGANIZATION: University of Illinois
Chicago, Illinois 60612-7205

REPORT DATE: October 1998

TYPE OF REPORT: Annual

PREPARED FOR: Commander
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	October 1998	Annual (1 Oct 97 - 30 Sep 98)

4. TITLE AND SUBTITLE

Use of Genetic Suppressor Elements to Identify Potential Breast Tumor Suppressor Genes

5. FUNDING NUMBERS

DAMD17-97-1-7293

6. AUTHOR(S)

Marija Zeremski

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

University of Illinois
Chicago, Illinois 60612-7205

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

U.S. Army Medical Research and
Materiel Command
504 Scott Street
Fort Detrick, Maryland 21702-5012

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for public release; distribution unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

ING1 is recently isolated breast cancer candidate tumor suppressor gene. In this study, structure and expression of the mouse *ing1* gene were analyzed. *Ing1* is transcribed from three differently regulated promoters. Each of the resulting transcripts represents two exons; they share a long common region encoded by a common exon and differ in their 5'-exon sequences encoded by alternative 5'-exons. Three alternative transcripts of *ing1* encode two proteins one of which is a truncated version of the other. Protein sequences encoded by mouse and human *ing1* are highly conserved. Expression of *ing1* mRNA expression correlates with cell proliferation suggesting a connection between cell growth regulation and *ing1* expression. Embryonic stem cells with heterozygous deletion of *ing1* common exon are prepared as a first step in the generation of *ing1* knockout mice. Analysis of *ing1* knockout animals should give important information about the role *ing1* might have in carcinogenesis.

14. SUBJECT TERMS	15. NUMBER OF PAGES
Breast Cancer	53
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Unlimited

FOREWORD

Opinions, interpretations, conclusions and recommendations are those of the author and are not necessarily endorsed by the U.S. Army.

____ Where copyrighted material is quoted, permission has been obtained to use such material.

____ Where material from documents designated for limited distribution is quoted, permission has been obtained to use the material.

____ Citations of commercial organizations and trade names in this report do not constitute an official Department of Army endorsement or approval of the products or services of these organizations.

X In conducting research using animals, the investigator(s) adhered to the "Guide for the Care and Use of Laboratory Animals," prepared by the Committee on Care and use of Laboratory Animals of the Institute of Laboratory Resources, national Research Council (NIH Publication No. 86-23, Revised 1985).

X For the protection of human subjects, the investigator(s) adhered to policies of applicable Federal Law 45 CFR 46.

X In conducting research utilizing recombinant DNA technology, the investigator(s) adhered to current guidelines promulgated by the National Institutes of Health.

X In the conduct of research utilizing recombinant DNA, the investigator(s) adhered to the NIH Guidelines for Research Involving Recombinant DNA Molecules.

NA In the conduct of research involving hazardous organisms, the investigator(s) adhered to the CDC-NIH Guide for Biosafety in Microbiological and Biomedical Laboratories.

Marija Zefren 10.22.1998.

PI - Signature Date

TABLE OF CONTENTS

	Page
Introduction	2
Body	3
Conclusions	4
References	6
Apendix	7

INTRODUCTION

Tumor suppressor genes are defined as elements whose loss or inactivation promotes neoplastic transformation (1,2). Mutations in tumor suppressor genes play a key role in both genetic predisposition of different forms of cancer and in the etiology of spontaneous malignancies. The goal of our research was the isolation and analysis of tumor suppressor genes whose lost or inactivation is involved in the formation or progression of breast cancer. To identify new tumor suppressor genes, we are using a recently developed genetic suppressor element (GSE) approach. This strategy is based on the expression selection of transforming GSEs, which promote transformation presumably by suppressing tumor suppressor genes from which they derived (3,4).

To isolate breast cancer tumor suppressor genes, we decided to screen GSE library specific for normal breast cells (4). This library is enriched for sequences derived from putative tumor suppressor genes, which is increasing the possibility of their isolation in different selection procedures. Previous screening of this library showed that it contains biologically active GSEs, since selection for the tumor formation in nude mice led to the isolation of several transforming elements. The strongest of the transforming GSEs was used for cloning its corresponding gene, named *ING1* (4). Expression of *ING1* is upregulated in senescent human fibroblasts (5), while ectopic expression of *ING1* cDNA leads to arrest in the G1 phase of the cell cycle or induces apoptosis in several cell types (6). *ING1* encodes nuclear protein $p33^{ING1}$ that cooperates with p53 in growth regulation by modulating the ability of p53 to act as a transcriptional activator (7). Limited analysis of tumor cell lines revealed reduced expression of *ING1* in several breast carcinoma cell lines (4), which indicates its involvement in the development of breast cancer. *ING1* is mapped to human chromosome 13q34 (8), and loss of heterozygosity in this chromosomal locus has been reported in squamous-cell carcinomas of head and neck. These properties of *ING1* allowed defining it as a candidate tumor suppressor gene, with high probability of involvement in the breast cancer formation or progression.

Goal of my research was the isolation of the new transforming GSEs from the GSE library specific for normal breast cells, as a first step towards the isolation of the new breast tumor suppressor genes. Since the screening of the GSE library is a long process, in parallel, I decided to study the *ING1* candidate tumor suppressor gene, since there is strong indications about its involvement in the formation of breast cancer. All the information regarding the $p33^{ING1}$ function was obtained so far only through *in vitro* experiments. However, in order to really understand a role that a certain gene might have, it is necessary to analyze its expression, regulation, and function in the living organism, to see its involvement in development and other physiological processes that can not be analyzed on isolated tissue culture cells. To do these kind of studies, it is necessary to have a model system that will allow easy and efficient genetic manipulation with the gene of interest. In order to do detailed analysis of *ING1*, I decided to use a mouse model, which is the ideal system for different kinds of genetic studies. As a first and essential step towards understanding its function and involvement in physiological processes as well as carcinogenesis, I analyzed the structure and regulation of the mouse *ing1* gene.

Cloning and structural analysis of the mouse *ING1* homologue was designed as a first step in the generation of *ing1* knockout animals. Those animals would provide valuable data on the role of the gene in normal development, but they could also give important insight into the role that *ING1* might have in carcinogenesis. The involvement of *ING1* in the p53 signaling pathway suggests that it could be a tumor suppressor by itself, in which case *ing1* knockout may lead to the consequences similar to those associated with *p53* deficiency in animals.

BODY

In order to isolate GSEs for the breast cancer tumor suppressor genes, I used GSE library prepared from cDNA specific for normal human breast epithelial cells mentioned in introduction (4). This library, cloned into pLNCX retroviral vector, is enriched for the sequences derived from tumor suppressor genes. Previous experiments showed that it contain biologically active GSEs, which is making it a good source for the isolation of genes whose inactivation is involved in breast cancer development. As test cells I used MCF10A (9), human pseudonormal immortalized breast cell line with low efficiency of growth in semi-solid media, low background in spontaneous focus formation and non-tumorigenic for nude mice (unpublished observation). Library was delivered to test cells using BING packaging cell line (10) for virus preparation. Infected cells were used in three different selection procedures: selection for foci formation, selection for colony formation in semi-solid agar, and selection for tumorigenicity in nude mice. In order to isolate biologically active GSEs several rounds of selection should be performed, and this work is still in progress.

In parallel, I studied the *ING1* candidate tumor suppressor gene, since there is strong indications about its involvement in the formation of breast cancer. In order to do detailed analysis of *ING1*, I used a mouse model, which is the ideal system for different kinds of genetic studies. The first step in this approach was the isolation of the mouse *ING1* homologue. In the previous year, I isolated mouse *ing1*, determined its structure, including precise mapping of the beginning and end points of *ing1* transcription as well as exon-intron structure, and analyzed its expression and regulation.

Mouse *ing1* is transcribed from three differently regulated promoters localized within a 4 kb segment of genomic DNA. Two of them are TATA-less promoters with a loose initiator sequence that includes a few tightly clustered transcription start sites and multiple Sp1 binding sites upstream from the initiator. The third promoter has TATA-like box, putative CAAT box and multiple Sp1 binding sites. Three alternative transcripts of *ing1* share a long common region encoded by a common exon and differ in their 5'-exon sequences encoded by alternative 5'-exons. Only one of the 5'-end exons contains a protein-coding sequence, while translation of the two other transcripts starts from the initiator codon within the common region. Three alternative transcripts of *ing1* encode two proteins ($p37^{ing1}$ and $p31^{ing1}$), one of which is a truncated version of the other lacking 104 N-terminal amino acids. Protein sequences of mouse and human *ing1* genes show a high degree of similarity, especially in their C-terminal regions that contain a PHD finger domain. At the nucleotide level, sequence similarity between human and mouse

cDNAs was high throughout the whole coding regions and much lower in the 3'-untranslated region. Several conserved regions localized within 3'-untranslated sequences of *ing1* possibly map some functional domains within *ing1* mRNA that could be involved in regulation of the mRNA transport, translation or stability.

Ing1 is expressed in adult mouse tissues at moderate levels, with the highest expression in the thymus and testis. Two isoforms are ubiquitously expressed with the highest levels in the thymus, while the third is shown to be testis specific. *Ing1* is expressed during the entire embryogenesis with the highest levels in the yolk sac of the early embryo and inner compartments of bones of the later stage embryos. Analysis of the cell lines showed that *ing1* expression correlates with cell proliferation suggesting a connection between cell growth regulation and *ing1* expression. Differences between expression patterns at mRNA and protein levels indicate possible regulation of *ing1* at the levels of translation or protein stability.

The work concerning structure and regulation of mouse *ing1* gene is submitted to Molecular and Cellular Biology, and I am sending a copy of the manuscript in Apendix. The manuscript includes detailed explanation about how this work is done, including experimental methods, results and discussion.

In order to make *ing1* deficient animals we decided to delete common exon of *ing1* in their genome. To do that, we applied widely used technique of positive/negative selection (11). Targeting vector, containing a neomycin resistance (*neo*) gene within homologous regions of the target sequence and a herpes simplex virus thymidine kinase (*TK*) gene outside of these regions, was made and introduced into embryonic stem (ES) cells. If site-specific recombination occurs in cells, integration of *neo* and elimination of *TK* gene will be observed. Selection of transfected ES cells to both G418 and ganciclovir resulted in isolation of cell population enriched with desired clones. Legitimate homologous recombination in these selected ES cells was confirmed by Southern blot hybridization, which revealed four colonies with heterozygous deletion within *ing1* gene. ES cells with a heterozygous deletion of the *ing1* common exon are currently being used for the generation of *ing1* knockout mice.

CONCLUSIONS

As a first step towards the better understanding of the mechanisms of function of *ing1* gene as well as its role in carcinogenesis, I analyzed its structure and regulation. I precisely mapped the beginning and end points of *ing1* transcription and determined exon-intron structure of the gene. The mouse *ing1* gene is transcribed from three differently regulated promoters. Three alternative transcripts of *ing1* share a long common region encoded by a common exon and differ in their 5'-exon sequences encoded by alternative 5'-exons. Only one of the 5'-exons contains protein-coding sequence while translation of two other transcripts starts from the initiator codon within the common region. Three alternative transcripts of *ing1* thus encode two different proteins, different only in size. Comparison of the mouse and human *Ing1* sequences reveals extremely high levels of homology.

The results of expression analysis indicated that mouse *ing1* is a subject of regulation both at the protein and mRNA levels and that different isoforms have different expression patterns. Expression of *ing1* mRNA correlates with cell proliferation suggesting a connection between cell growth regulation and *ing1* expression.

First step in the generation of *ing1* knockout mice, preparation of ES cells with a heterozygous deletion of the *ing1* common exon, is finished. Completion of this project should give valuable information regarding the possible role of this gene in carcinogenesis.

REFERENCES

1. Weinberg, R.A. (1991). Tumor suppressor genes. Science 254, 1138-1146.
2. Levine, A.J. (1993). The tumor suppressor genes. Ann. Rev. Biochem. 62, 623-651.
3. Ossovskaya, V.S., Mazo, I.A., Chernov, M.V., Chernova, O.B., Strezoska, Z., Kondratov, R., Stark, G.R., Chumakov, P.M., and Gudkov, A.V. (1996). Dissection of p53 functions by genetic suppressor elements: distinct biological effects of separate p53 domains. Proc. Natl. Acad. Sci. USA 93, 10309- 10314.
4. Garkavtsev, I.A., Kazarov, A.R., Gudkov, A.V., and Riabowol, K. (1996). Suppresion of the novel growth inhibitor $p33^{ING1}$ promotes neoplastic transformation.Nature Genetics 14, 415-420.
5. Garkavtsev, I.A., and Riabowol, K. (1997). Extension of the replicative life span of human diploid fibroblasts by inhibition of the p33ING1 candidate tumor suppressor. Mol. Cel. Biol. 17, 2014-2019.
6. Helbing, C.C., Veillette, C., Riabowol, K., Johnston, R.N., and Garkavtsev, I. (1997). A novel candidate tumor suppressor, *ING1*, is involved in the regulation of apoptosis. Cancer Res. 57, 1255-1258.
7. Garkavtsev, I., Grigorian, I.A., Ossovskaya, V.S., Chernov, M.V., Chumakov, P.M., and Gudkov, A.V. (1998). The candidate tumor suppressor $p33^{ING1}$ cooperates with p53 in cell growth control. Nature 391, 295-298.
8. Zeremski, M., Horrigan, S.K., Grigorian, I.A., Westbrook, C.A., and Gudkov, A.V. (1997). Localization of the candidate tumor suppressor gene *ING1* to human chromosome 13q34. Somat. Cell Mol. Genet. 23, 233-236.
9. MCF10A cell line: ATCC Cell Lines and Hybridomas eight edition 1994. Under CRL-10317.
10. Pear, W.S., Nolan, G.P., Scott, M.L., and Baltimore, D. (1993). Production of high titer helper-free retroviruses by transient transfection, Proc. Natl. Acad. Sci. USA 90, 8392-8396.
11. Mansour, S.L., Thomas, K.R., and Capecchi, M.R., (1988). Disruption of the protooncogene int-2 in mouse embryo-derived stem cells: a general strategy for targeting mutations to non-selectable genes. Nature 336, 348-352.

APENDIX

I am including in my report manuscript of the paper submitted to Molecular and Cellular Biology, which is explaining in details results of the study addressing questions of *ing1* structure and regulation.

Structure and Regulation of the Mouse *ing1* Gene: Three Alternative Transcripts Encode Two PHD Finger Proteins Differing in Their N-termini

MARIJA ZEREMSKI[1], JASON E. HILL[2], IGOR V. GARKAVTSEV[3], LUDA DIATCHENKO[2], EUGENE V. KOONIN[4] AND ANDREI V. GUDKOV[1]*

Department of Molecular Genetics, College of Medicine, University of Illinois at Chicago, Chicago, IL 60607, USA[1], CLONTECH Laboratories, Inc., Palo Alto, CA 94303[2], Department of Functional Genomics, Genome Therapeutics Corp., Waltham MA 02154[3], and National Center for Biotechnology Information (NCBI), National Library of Medicine, National Institutes of Health, Bethesda, Maryland 20894[4].

*Corresponding author. Mailing address: Dept. of Molecular Genetics (M/C 669), University of Illinois at Chicago, 900 South Ashland Avenue, Chicago, IL 60607. Phone: (312) 413-0349. Fax: (312) 996-0683. Electronic mail address: gudkov@uic.edu

Keywords: candidate tumor suppressor gene, alternative splicing

Running title: Mouse *ing1* gene

ABSTRACT

Structure and expression of the mouse homologue for the human *ING1* candidate tumor suppressor gene were analyzed. Mouse *ing1* is transcribed from three differently regulated promoters localized within a 4 kb region of genomic DNA. Each of the resulting transcripts represent two exons; they share a long common region encoded by a common exon and differ in their 5'-exon sequences encoded by alternative 5'-exons. Only one of the 5'-exons contains protein-coding sequence; translation of two other transcripts starts from the initiator codon within the common region. Three alternative transcripts of *ing1* thus encode two proteins ($p37^{ing1}$ and $p31^{ing1}$), one of which is a truncated version of the other lacking 104 N-terminal amino acids. Protein sequences encoded by mouse and human *ing1* are highly conserved, especially in their C-terminal regions that contain a PHD finger domain. $p33^{ING1}$ belongs to a small family of proteins from humans and yeast that have approximately the same size, show highly significant similarity to one another within the PHD finger domain and also contain an additional, N-terminal region with subtle but reliably detectable sequence conservation. Expression of *ing1* mRNA correlates with cell proliferation suggesting a connection between cell growth regulation and *ing1* expression.

INTRODUCTION

The *ING1* candidate tumor suppressor was identified as a result of a functional screening of genes, the suppression of which is associated with neoplastic transformation (5). Inhibition of *ING1* expression by antisense RNA promotes anchorage independent growth in mouse breast epithelial cells, increases the frequency of focus formation in NIH 3T3 cells, and prolongs the life span of diploid human fibroblasts in culture. *ING1* expression is upregulated in senescent human fibroblasts (6), and ectopic expression of *ING1* cDNA leads to G1 arrest or promotes apoptosis in several cell types (12). The accumulated observations indicating *ING1* participation in the negative regulation of cell proliferation, control of cellular aging, and apoptosis have defined *ING1* as a candidate tumor suppressor gene. *ING1* was mapped to human chromosome 13q34 (7, 26), and loss of heterozygosity in this chromosomal locus has been reported in squamous cell carcinomas of the head and neck (13).

We have recently found that the biological effects of *ING1* and *p53* are interrelated and require the activity of both genes: neither of the two genes can, on its own, cause growth inhibition when the other one is suppressed (9). Furthermore, activation of transcription from the *p21/WAF1* promoter, a key mechanism of p53-mediated growth control, depends on the expression of *ING1*. A physical association between $p33^{ING1}$ and p53 proteins is detected by immunoprecipitation. These results indicated that $p33^{ING1}$ is a component of the p53 signaling pathway, and that $p33^{ING1}$ cooperates with p53 in negative regulation of cell proliferation by modulating p53-dependent transcriptional activation.

In spite of the apparent importance of *ING1* in the control of cell proliferation, our knowledge of expression, regulation, and function of this gene remains incomplete. Moreover, GenBank contains two versions of *ING1* mRNA sequences differing in their 5'-ends; the origin of these differences is unknown and requires explanation. So far, all of the information about the function of $p33^{ING1}$ was obtained from *in vitro* experiments that involved ectopic expression of *ING1* cDNA or its suppression by antisense RNA. The analysis of the structure of the *ING1* gene and its regulation *in vivo* is an essential step towards the understanding of its function and involvement in developmental and physiological processes. This is particularly important due to the cooperation between *ING1* and *p53*, which suggests that the functioning of the p53 signaling pathway could be dependent on the regulation of *ING1* expression.

In the present work, we used the mouse orthologue of human the *ING1* gene for detailed structural and expression studies, with the goal of subsequently utilizing it as a system for extensive genetic analysis. *Ing1* was found to be a highly evolutionarily conserved gene with complex regulation which involves generation of alternative transcripts initiated from different promoters and translated into proteins that differ in structure and expression patterns.

MATERIALS AND METHODS

Animals

Organs and embryos of FVB/N mice were used for the RNA isolation and preparation of whole-body sections of embryos using cryostatic microtome for histo-blot hybridization.

Cell lines

Cultures of mouse embryonic fibroblasts were obtained from 10 day-old embryos. 10(1) cell line, a derivative of Balb 3T3 cells that spontaneously deleted both p53 alleles (11), was kindly provided by Arnold Levine. Pseudonormal mouse mammary gland epithelial cell line (NMuMG) was obtained from the ATCC collection. Ecotropic retroviral packaging cell line BOSC23 (16) was kindly provided by Warren Pear and David Baltimore (Massachusetts Institute of Technology). All cells were maintained in Dulbecco's modified Eagle medium containing 10% fetal bovine serum supplemented with penicillin and streptomycin. For the serum starvation experiments, 10(1) and NMuMG cells were kept with 0.5% fetal bovine serum for 48 hours (NMuMG cells) or 36 hours (10(1) cells). For the contact inhibition experiments, cells cultures were used 48 hours (NmuMG) or 36 hours (10(1)) after they became confluent. NMuMG and 10(1) cells were irradiated with 10 Gy of γ-radiation and used 24 hours after treatment. Mouse embryo fibroblasts were propagated until they reached senescence.

Plasmids and libraries

Retroviral vector pLXSN, used for the introduction of mouse and human *ING1* cDNA in NMuMG and 10(1) cells, was provided by A. Dusty Miller (14). A cDNA library from senescent mouse embryonic fibroblasts was constructed using SuperScript system (Gibco BRL) according to the manufacturer's protocol, and cloned into the lambda phage λ gt22A vector. Stratagene's 129SVJ Mouse Genomic Library cloned into the

Lambda FIX II vector was used for the isolation of the mouse genomic, *ing1* containing, clones.

Hybridization screening of cDNA and genomic libraries

cDNA and genomic library screenings were done according to the standard protocols (18) using ^{32}P-labeled human and mouse *ing1* cDNA probes, respectively.

5' and 3' RACE

Alternative 5' ends of the mouse *ing1* were isolated from mouse spleen and mouse brain cDNAs using the Marathon-ready cDNA kit (CLONTECH Laboratories), according to the protocol suggested by the manufacturer. AP1 adaptor-specific sense primer, provided by CLONTECH, and the *ing1*-specific antisense primer (5' - CCATCTGACTCACGATCTGGATCTTC - 3') were used for PCR. Nested PCR was performed using AP2 adaptor-specific sense primer, provided by CLONTECH, and *ing1*-specific antisense primer (5' - CTGCGGATCAGGGCCCTCTGGATGC - 3'). Precise determination of the 5' and 3' ends of the mouse *ing1* transcripts was done using the Marathon-2 cDNA Amplification kit based on the new SMART PCR cDNA synthesis technology (CLONTECH). It is based on the recently identified ability of M-MLV reverse transcriptase to add several nucleotides to the 3'-terminus of first-strand cDNA during the RT reaction (3). Briefly, when reverse transcriptase reaches the 5'end of the mRNA, it switches templates and continues synthesizing the SMART template-switching (TS) oligonucleotide. The resulting single-stranded cDNA contains the complete 5'end of the

mRNA as well as the sequence complementary to the TS oligonucleotide and is then selectively amplified by PCR. In these experiments, poly(A) RNA preparations isolated from the thymus and testis were used. The following sequences were used for the syntheis of antisense *ing1*-specific primers: 5' - AGGTGTGGTGGGATCGGCAACGC - 3' (for is*oform 1a), 5' - CGCGGGGAGCCAGAGCAGAGAAGGT - 3' (isoform 1c), and 5' - GGCGTGGCCTGTCATTGTCGCTG - 3' (isoform 1b). *Ing1*-specific sense primer 5' - GCGTGCTTCTTGCTACCAT - 3' was used for the PCR amplification of the 3' end.

Sequence analysis

Sequencing was done using a Sequenase Version 2.0 DNA Sequencing Kit (USB), or by the University of Chicago Cancer Research Center DNA Sequencing Facility. In all cases, both strands were read using multiple vector-specific or gene-specific primers. Protein sequence database searches were performed using the gapped BLASTP program and the PSI-BLAST program that iterates the search using profiles constructed from BLAST hits as queries for subsequent iterations (2). Multiple sequence alignments were constructed using the Gibbs sampling option of the MACAW program (15, 19).

Southern, northern, and western analyses were done according to standard protocols. IgG1 mouse monoclonal antibody against human recombinant $p33^{ING1}$ (8) was used for the detection of *ing1*-encoded proteins.

In situ hybridization ("histo-blotting")

Whole-body sections, 20 μm thick, were prepared from frozen embryos embedded in blocks of Tissue-Tek O.C.T. and stored at -70°C. Sections were placed on nitrocellulose (Schleicher&Schuell) to prepare "histo-blots" as previously described (10). Histo-blots were hybridized with ^{33}P-labeled RNA probes synthesized using Ambion's MAXIscript *In vitro* Transcription Kit. *Ing1* specific sense and antisense probes were synthesized on the pBLUESCRIPT plasmid with the fragment of *ing1* cDNA corresponding to the PHD finger domain using T3 polymerase for the antisense and T7 polymerase for the sense probe. β-actin antisense RNA was synthesized using the template provided by Ambion. Histo-blots were incubated in pre-hybridization solution (0.75 mg/ml yeast tRNA, 0.75 mg/ml poly(A), 50% formamide, 0.3M Tris pH 8.0, 1mM EDTA, 5x Denhardt's solution, 10% Dextran sulfate, 10mM DTT) at 42°C for 1-4 hours. Hybridization was carried out for 12-24 hours at 42°C in the same solution containing 1-5 x 10^7cpm of probe per ml. After hybridization, histoblots were washed in 4x SSC at room temperature for 20 min, in 2x SSC, 0.04 μg/ml RNase A at 37°C for 30 min, in 2x SSC at 37°C for 30 min, in 1x SSC at 60°C for 15-30 min and finally in 0.1x SSC at 60°C for 15-30 min. and exposed to X-ray film for 2-7 days.

RESULTS

Multiple transcripts of the mouse ing1 *gene differ in their 5'-end sequences*

In order to isolate the mouse orthologue of the *ING1* gene, we screened a cDNA library prepared from senescent mouse embryonic fibroblasts, using human *ING1* as a probe. The choice of the library was determined by the fact that *ING1* is expressed at

higher levels in senescent compared to normal human fibroblasts (6). As a result, several clones were isolated, and the longest two were sequenced. The clones were identical to each other and highly similar to human *ING1* through most of their length except for the 5' ends which were different and not homologous to the human gene (Figure 1). This observation could be an indication of alternative splicing of mouse *ing1*, but it could also be potentially explained by cloning artifacts occurring during the cDNA library preparation. To determine the structure of 5'-end sequences of *ing1,* we used a 5'-RACE technique for the isolation of cDNA sequences corresponding to the 5'-termini of *ing1* mRNA. cDNA was synthesized from mouse spleen and brain mRNA, and ligated to synthetic adaptors. PCR was performed on the cDNA using a sense primer specific for the adaptor and an antisense primer specific for the common part of *ing1* located close to the divergent region. Two fragments of different size were obtained and cloned, sequenced, and compared to the clones isolated from the cDNA library. The results of this comparison are schematically presented in Figure 1. One of the 5'-RACE products was identical to one of the previously isolated cDNA clones. Another product revealed the third variant of *ing1* cDNA, which again consisted of common and unique regions with the junction located exactly at the same position as in the other sequences. Figure 1 shows schematic alignments of the isolated mouse *ing1* cDNA clones and the 5' RACE products. All variants are identical to each other (and homologous to human *ING1*) up to the same nucleotide, and different only in their 5' ends.

In order to verify that the obtained cDNA structure reflects naturally existing transcripts, we analyzed *ing1* mRNA species by Northern hybridization using probes representing common or specific regions of isolated cDNAs. As shown in Figure 2, the

probe for the common region revealed multiple transcripts in RNA isolated from mouse liver, heart, and testis. Probes specific for individual transcripts, however, showed more simple hybridization patterns, which in combination covered the whole set of transcripts found by hybridization with the common probe. These results indicated that the cloned sequences represent the majority of multiple transcripts of *ing1* synthesized in the thymus and spleen.

Mapping of coding regions of ing1 *in mouse genomic DNA*

Comparison of sequences of individual cDNA clones showed differences in their 5'-regions suggesting that *ing1* has multiple alternative 5'-exons (Fig. 1). To verify this hypothesis, we determined the structure of the mouse *ing1* gene. We isolated phage clones carrying sequences homologous to the mouse *ing1* cDNA by hybridization screening of a mouse genomic library. These clones were mapped by restriction digestion analysis, in combination with Southern blot hybridization, with the probes corresponding to the different *ing1* parts. The interpretation of the obtained results is shown in Figure 3.

Southern blot hybridization analysis of *ing1*-related sequences in the mouse genome indicated that ing1 is a single gene with no obvious close family members. Comparison of genomic and cDNA sequences of *ing1* using PCR with different *ing1* specific primers (data not shown), as well as with the sequencing data, showed that most of the transcribed sequences of the *ing1* gene come from a single exon. Alternative 5' ends are encoded by different exons positioned upstream from the common exon. Comparison of the sequences of cDNA and genomic clones, as well as 5' RACE

products, revealed three isoforms of mouse *ing1* that differ from each other only in their 5' ends, which also indicates that each isoform is most probably expressed from its own promoter.

The length of the isolated cDNA clones appeared to be significantly shorter than that of the mRNA species detected by Northern hybridization (Fig. 2), suggesting that part of the transcribed sequences were missing from the isolated cDNAs. To determine the exact start sites of *ing1* transcription, we used a new procedure called "SMART-based 5'RACE" as described in Materials and Methods. Using this method, we were able to extend the cDNA sequences for the 5'end of isoform 1b. This exon contains an extremely G/C rich region that probably resulted in the incomplete sequence obtained from the original 5'-RACE (Figure 4).

Comparing the 3'ends of human and mouse *ING1* cDNAs showed that although the mouse transcript was flanked by a poly(A) stretch, its 3'-untranslated region was significantly shorter than the human one. Analysis of the genomic clone with the mouse *ing1* sequence revealed the presence of a long poly(A) stretch that could potentially be used as a primer-binding sequence for reverse transcription initiated from oligo(dT) primers. Moreover, the alignment with the 3' end of the human *ING1* cDNA sequence resumes downstream from this genomic polyA stretch. The exact end of *ing1* transcription in mouse was determined using the same method applied for the generation of 5'-sequences. Mouse *ing1* cDNA was in fact found to be longer than was originally thought (Figure 4). Sequence comparison showed that mouse and human *ING1* transcripts terminate at the same point, and share significant levels of similarity up to the very 3' ends.

Sequence analysis of ing1

Nucleotide and predicted amino acid sequences of the three mouse *ing1* cDNA isoforms and the promoter regions for all three alternative transcripts are shown in Figures 5 and 6. It was found that the 1a and 1c *ing1* isoforms have several tightly clustered transcription start sites. Sequences upstream of the initiation sites (putative promoters) lack TATA boxes, but they do contain a sequence corresponding to the loose initiator (Inr) consensus PyPyANT/APyPy that includes the transcription start sites (Figure 5). In both cases, areas upstream from the Inr are extremely GC-rich with multiple Sp1-binding sites. All of these features are typical for TATA-less promoters (21, 22). Binding sites of inducible factors that are usually present further upstream in the promoter area (e.g. CAAT box, Oct-1 and Oct-2 binding site, NFκB or ATF binding site) were not found in these promoters. Transcription of the *ing1* isoform 1b starts approximately 30 nucleotides downstream from the TATA-like box. A putative CAAT box is present 60 nucleotides upstream from the TATA-like sequence. Since the transcription initiation sites for the 1a and 1b isoforms are only about 200 bp apart, there is a possibility that their promoters share some of the regulatory sequences including, for example, a number of Sp1 binding sites located upstream from this area.

All three alternative transcripts of mouse *ing1* contain the same long open reading frame, although the sizes of the predicted proteins are different. While isoform 1b encodes a protein of 279 amino acids, the other two isoforms are predicted to encode a shorter protein product of 175 amino acids, which lacks 104 N-terminal amino acids. Translation of these two products is expected to start from the initiation codon, which is

located in the beginning of the common exon. Alignment of the predicted amino acid sequences with that of human $p33^{ING1}$ (GenBank accession number AF044076) revealed high similarity between the mouse and human proteins (89% sequence identity) (Fig. 7a). At the nucleotide level, sequence similarity between the human and mouse cDNAs was high throughout the whole coding regions and much lower in the 3'-untranslated region that contained, however, several conserved stretches of nucleotides (Fig. 7b).

The non-redundant protein sequence database at NCBI was searched using the *ing1*-encoded protein sequence as a query and a highly significant sequence similarity was detected with a human paralogue of $p33^{ING1}$, three uncharacterized proteins from the budding yeast *Saccharomyces cerevisiae*\ and their homologue from fission yeast *Schizosaccharomyces pombe* (probability of the similarity being observed by chance $<10^{-12}$). All of these proteins have approximately the same size and contain a C-terminal PHD finger domain (1, 20). The sequence conservation in the PHD domain in these six protein sequences is striking, with 13 invariant residues, in addition to the 8 metal-chelating cysteines and histidines that are conserved in all PHD fingers (Fig. 8). Multiple alignment analysis resulted in the delineation of an additional N-terminal region that is conserved in these proteins (Fig. 8). This region consists of approximately 100 amino acid residues, includes two distinct conserved motifs (Fig. 8), and shows a fairly subtle similarity that was not statistically significant in the context of the screening of the complete database, except for the conservation between the two human paralogs. However, in the reduced search space defined by the presence of the PHD finger, it was

shown that the probability of finding this level of similarity by chance was below 10^{-13} for the distal motif and below 10^{-5} for the proximal motif.

Expression of *ing1* transcripts in adult and embryonic mouse tissues

We analyzed the expression of *ing1* mRNA in the organs of adult mice and mouse embryos at different stages of development by Northern blot hybridization. The probe representing the common exon of the *ing1* gene revealed multiple mRNA size classes that represent alternative transcripts of *ing1* (Figure 9). The overall *ing1* mRNA expression is most abundant in thymus and testis; much lower levels were detected in the rest of the tissues tested that also differed in the content and relative intensity of the hybridizing bands. The same pattern of *ing1* mRNA expression was observed in p53-deficient mice showing no direct effect of p53 on *ing1* regulation (data not shown). In embryos, the highest expression was found on the 11th day of development and was characterized by changes in the relative expression of different classes of mRNA transcripts.

Patterns of *ing1* expression in embryogenesis were also checked by *in situ* hybridization. We used the histoblotting technique, in which embryonic sections were fixed on a nitrocellulose membrane, and then hybridized with the RNA probes corresponding to the common part of the *ing1* cDNA. Besides sense and antisense *ing1* specific RNA probes, we also used an antisense probe for β-actin, a gene with ubiquitous expression, as an internal standard. Sections were prepared from 10-, 12-, 16- and 18-day embryos. Results of *in situ* hybridization are shown in Figure 10. *Ing1* is uniformly expressed in the whole mouse embryo at all stages of development examined.

This is in agreement with the result obtained with the adult mouse tissues, where *ing1* is expressed ubiquitously and at similar levels in all analyzed organs. However, at embryonic day 10 higher expression levels were observed in the yolk sac, while at day 16 and 18 of development, higher levels of expression were detected in inner compartments of bones and probably match areas of ongoing ossification.

In order to determine patterns of expression of different *ing1* transcripts in mouse tissues and during embryogenesis, Northern blots were hybridized with probes corresponding to the alternative starts of *ing1* (Figure 9). The 1b and 1c variants were expressed ubiquitously in all adult mouse tissues tested, with the highest levels in thymus. Both variants were expressed in the embryos at all stages of development analyzed, with the highest levels at day 7 in the case of 1c, or day 11 in the case of the 1b variant. Out of all adult tissues analyzed, mRNA for isoform 1a was expressed only in testis; it was also seen in the 11-day embryo. While there is no detectable signal with the 1a specific probe in the 7 day embryo, traces of hybridization could be detected in mRNA from day 15 and day 17 embryos, which indicates an extremely low expression level at these points of mouse embryogenesis.

Expression of the mouse *ing1* gene in cell lines depends on cell proliferation

To check whether *ing1* expression was cell cycle dependent, we analyzed RNA from two mouse cell lines, NMuMG and 10(1), at different stages of the cell cycle by Northern blot hybridization using total *ing1* cDNA as a probe. *Ing1* expression was also compared in senescent versus young, dividing mouse embryonic fibroblasts. Results are shown in Figure 11. In both cell lines, *ing1* was expressed at higher levels in dividing

compared to quiescent cells (quiescence is induced either by serum starvation, contact inhibition, or γ-irradiation). However, in 10(1) cells the difference in expression was specific only for the upper *ing1* specific band, which corresponds to the 1b variant. In mouse embryonic fibroblasts *ing1* was expressed at very low levels without any differences between dividing and senescent cells.

Ing1 expression in cell lines was also analyzed by Western blotting using an IgG1 mouse monoclonal antibody produced against human recombinant $p33^{ING1}$ (8). This antibody has previously been shown to specifically detect the denatured form of mouse $p33^{ING1}$ protein in Western immunoblot protocols. Cell lysates were produced from dividing, contact inhibited, serum starved, or γ-irradiated 10(1) and NMuMG cells. Also, as a control, cell lysates were prepared from BOSC23 cells transfected with mouse the 1b *ing1* variant or with the human truncated form of *ING1* that should give the protein product of the same size as mouse 1c and 1a variants, respectively (Figure 12). In both cell lines, two protein products were detected, 31kD and 38kD in size, which correspond to the truncated *ING1* and 1b *ing1* protein products respectively. The 38kD protein was present at the same levels in dividing and quiescent cells, while the 33kD protein was present at higher levels in dividing compared to non-dividing cells. The variation of expression of *ing1* differs significantly at RNA and protein levels, indicating that this gene could be regulated at the level of translation or protein stability.

DISCUSSION

ING1 was originally described as a gene whose suppression promotes neoplastic transformation (5). Consistent with that, ectopic expression of the first isolated human cDNA was growth suppressive for different cell lines. *ING1* was found to encode a

nuclear protein termed $p33^{ING1}$ that was shown to functionally and physically interact with p53 in cell growth regulation (7, 9). Now, after we have characterized the structure and expression of the mouse *ing1* gene, it became clear that *ing1* regulation is more complicated than was originally thought. *Ing1* is transcribed from at least three differently regulated promoters and the resulting transcripts encode at least two different proteins. All the transcripts share a common region encoded by a common exon but differ in their 5'-exons. Two of these alternative exons do not contain protein-coding sequences (isoforms 1a and 1c) while the third one does (isoform 1b). Consistently, one of the *ing1* transcripts encodes a 37 kD protein ($p37^{ING1}$) while two others are translated into a shorter protein of 24 kD that surprisingly runs as if it was 31 kD ($p31^{ING1}$). Structures of mouse and human *ing1* genes are likely to be similar considering the high degree of evolutionary conservation of *ing1* sequences and the fact that the two versions of human *ING1* that are currently available in GenBank also share a large common part but have different N-termini. It is noteworthy that one of the human variants of *ING1* (GenBank accession number AF001954) does not have a homologue with similar 5'-sequences among the identified mouse *ing1* isoforms which raises the possibility that there could be more variants of mouse *ing1* transcripts that so far have not been isolated.

The conserved regions in the 3'-untranslated sequences of *ing1* (identified by aligning the mouse and human cDNAs) may potentially be involved in the regulation of mRNA transport, translation, or stability. Remarkably, one of these regions overlaps with the sequence of the anti-*ING1* genetic suppressor element (GSE) that is known to be effective against both human and mouse genes (5). This means that the GSE targets a specific region of the mRNA that is significantly shorter than the whole GSE sequence.

$p33^{ING1}$ and its homologs in such a distantly related species as yeast contains a remarkably conserved PHD finger domain and an additional, weakly conserved domain of unknown function. PHD fingers have been shown to bind DNA (20, 23) whereas to our knowledge, there is no evidence that they mediate protein-protein interactions. Therefore, it seems likely that the C-terminal PHD domain in the $p33^{ING}$ family of proteins is involved in specific DNA binding which may be important for transcription regulation. Given the outstanding conservation of this domain across the large phylogenetic distance that separates humans and yeast, the specificity of DNA binding with respect to particular binding sites is expected to be conserved either in terms of DNA sequence, distinct features of chromatin structure, or both. It may be further surmised that the conserved N-terminal domain is involved in specific protein-protein interactions that couple transcription regulation by $p33^{ING1}$ and its homologs to other elements of cell cycle control. The conservation of domain organization in $p33^{ING1}$ and its yeast homologs is particularly notable given that yeast does not encode any homologs of p53. Thus it appears that $p33^{ING1}$ is an ancient cell cycle regulator whose interaction with p53 is a later evolutionary addition.

The results of expression analysis indicated that mouse *ing1* is a subject of regulation both at the protein and mRNA levels and that different isoforms have different expression patterns. Thus, $p37^{ING1}$ is ubiquitously expressed in all tissues analyzed with elevated mRNA expression levels in the thymus, while levels of the $p31^{ING1}$ protein vary dramatically among organs and cell lines.

Sizes of the *ing1* mRNA transcripts detected by Northern hybridization correlate well with the length of cDNA sequences isolated from all the tissues except testis.

Probes specific for each of the identified splice variants reveal multiple RNA bands indicating a variability of the *ing1* transcripts in this organ. It is likely that in testis *ing1* mRNAs are either terminated or processed differently from the rest of the tissues.

There is a clear link between the proliferation rate of the cell and *ing1* expression both at the protein and mRNA levels. This observation may reflect cell cycle dependence of *ing1* transcription detected earlier for the human *ING1* gene; the biological significance of this regulation is not obvious so far. We failed to observe up-regulation of mouse *ing1* in senescent cells as it has been reported for human *ING1* (6) which suggests that regulation patterns of mouse and human *ing1* genes may not always be the same.

The complex regulation of *ing1* points to possible functional differences between the two protein isoforms. In fact, a similar type of regulation (alternative initiation leading to variability of 5'-exons) found in other tumor suppressor genes, namely *BRCA1*, *APC,* and *INK4* (4, 17, 24, 25), is associated with the generation of proteins with different functions (in the case of *INK4,* different reading frames are translated in the alternative transcripts). Experiments on functional characterization of human $p38^{ING1}$ and $p31^{ING1}$ showed significant differences in the biological activity of these two isoforms of the *ing1*-encoded protein products (Grigorian et al., in preparation).

One of the key questions about *ING1* that remains unanswered is whether the deregulation of this gene is involved in naturally occurring cancer. The existence of at least two differentially regulated protein products encoded by *ING1* with potentially different functions clearly means that any attempts to investigate the potential cancer

relevance of *ING1* should involve separate analysis of the alternative isoforms of this gene.

ACKNOWLEDGMENTS

We thank Ilya Mazo for his mouse cDNA library and Elena Komarova for providing histoblots of mouse embryos for *in situ* hybridization. This work was supported by grants from the National Institutes of Health CA60730 and CA75179 to A.V.G. and by grant from the Department of the Army DAMD17-97-1-7293 to M.Z.

REFERENCES

1. **Aasland, R., T. J. Gibson, and A. F. Stewart.** 1995. The PHD finger: implications for chromatin-mediated transcriptional regulation. Trends Biochem. Sci. **20**:56-59.
2. **Altschul, S. F., T. L. Madden, A. A. Schaffer, J. Zhang, Z. Zhang, W. Miller, and D. J. Lipman.** 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. **25**:3389-3402
3. **Chenchik, A., Y. Y. Zhu, L. Diatchenko, R. Li, J. Hill, and P.D. Siebert.** 1998. Generation and use of high-quality cDNA from small amounts of total RNA by SMART^{TM} PCR. In Gene cloning and analysis by RT-PCR, pp. 305-319. (ed. P. D. Siebert and J. W. Larrick), BioTechniques Books.
4. **Cui, J. Q., H. Wang, E. S. Reddy, and V. N. Rao.** 1998. Differential transcriptional activation by the N-terminal region of *BRCA1* splice variants *BRCA1a* and *BRCA1b*. Oncol. Rep. **5**:585-589.
5. **Garkavtsev, I., A. Kazarov, A. Gudkov, and K. Riabowol.** 1996. Supression of the novel growth inhibitor p33^{ING1} promotes neoplastic transformation. Nat. Genet. **14**:415-420.
6. **Garkavtsev, I., and K. Riabowol.** 1997. Extension of the replicative life span of hyman diploid fibroblasts by inhibition of the p33^{ING1} candidate tumor suppressor. Mol. Cell. Biol. **17**:2014-2019.
7. **Garkavtsev, I., D. Demetrick, and K. Riabowol.** 1997. Cellular localization and chromosome mapping of a novel candidate tumor suppressor gene. Cytogenet. Cell Genet. **76**:176-178.

8. **Garkavtsev, I., D. Boland, J. Mai, H. Wilson, C. Veillette, and K. Riabowol.** 1997. Specific monoclonal antibody raised against the $p33^{ING1}$ tumor suppressor. Hybridoma **16**:537-540.

9. **Garkavtsev, I., I. A. Grigorian, V. S. Ossovskaya, M. V. Chernov, P. M. Chumakov, and A. V. Gudkov.** 1998. The candidate tumor suppressor $p33^{ING1}$ cooperates with p53 in cell growth control. Nature **391**:295-298.

10. **Gudkov, A. V., K. N. Kashkin, T. E. Zaitsevskaya, and S. M. Troyanovsky.** 1989. Histo-blotting: hybridization in situ detection of specific RNAs on tissue sections transferred on nitrocellulose. Int. J. Cancer **44**:1052-1056.

11. **Harvey, D. M. and A. J. Levine.** 1991. p53 alteration is a common event in the spontaneous immortalization of primary BALB/c murine embryo fibroblasts. Gen. Dev. **5**:2375-2385.

12. **Helbing, C. C., C. Veillette, K. Riabowol, R. N. Johnston, and I. Garkavtsev.** 1997. A novel candidate tumor suppressor, *ING1*, is involved in the regulation of apoptosis. Cancer Res. **57**:1255-1258.

13. **Maestro, R., S. Piccinin, C. Doglioni, D. Gasparotta, T. Vukosavljevic, S. Sulfaro, L. Barzan, and M. Boiocchi.** 1996. Chromosome 13q deletion mapping in head and neck squamous cell carcinomas: identification of two distinct regions of preferential loss. Cancer Res. **56**:1146-1150.

14. **Miller, A. D., and G. J. Rosman.** 1989. Improved retroviral vectors for gene transfer and expression. BioTechniques **7**:980-986.

15. **Neuwald, A. F., J. S. Liu, and C. E. Lawrence.** 1995. Gibbs motif sampling: detection of bacterial outer membrane protein repeats. Protein Sci. **8**:1618-1632.

16. **Pear, W. S., G. P. Nolan, M. L. Scott, and D. Baltimore.** 1993. Production of high titer helper-free retroviruses by transient transfection. Proc. Natl. Acad. Sci. USA **90**:8392-8396.

17. **Quelle, D. E., F. Zindy, R. A. Ashmun, and C. J. Sherr.** 1995. Alternative reading frames of the INK4a tumor suppressor gene encode two unrelated proteins capable of inducing cell cycle arrest. Cell **83**:993-1000.

18. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

19. **Schuler, G. D., S. F. Altschul, and D. J. Lipman.** 1991. A workbench for multiple alignment construction and analysis. Proteins **9**:180-190

20. **Schindler, U., H. Beckmann , and A. R.Cashmore.** 1993. HAT3.1, a novel Arabidopsis homeodomain protein containing a conserved cysteine-rich region. Plant J. **4**:137-150

21. **Smale, S. T.** 1994. Core promoter architecture for eukaryotic protein-coding genes. In Transcription: Mechanisms and regulation, pp. 63-80 (ed. R. C. Conaway and J. W. Conaway), Raven Press, New York, NY.

22. **Smale, S. T.** 1997. Transcription initiation from TATA-less promoters within eukaryotic protein-coding genes. Biochim. Biophys. Acta **1351**:73-88.

23. **Stassen, M. J., D. Bailey, S. Nelson, V. Chinwalla, and P. J. Harte.** 1995. The Drosophila trithorax proteins contain a novel variant of the nuclear receptor type DNA binding domain and an ancient conserved motif found in other chromosomal proteins. Mech Dev **52**:209-223.

24. **Thliveris, A., W. Samowitz, N. Matsunami, J. Groden, and R. White.** 1994. Demonstration of promoter activity and alternative splicing in the region 5' to exon 1 of the *APC* gene. Cancer res. **54**:2991-2995.

25. **Xu, C.-F., M. A. Brown, J. A. Chambers, B. Griffiths, H. Nicolai, and E. Solomon.** 1995. Distinct transcription start sites generate two forms of *BRCA1* mRNA. Hum. Mol. Genet. **4**:2259-2264.

26. **Zeremski, M., S. Horrigan, I. A. Gregorian, C. Westbrook, and A. V. Gudkov.** 1997. Chromosomal localization of $p33^{ING1}$ candidate tumor suppressor gene. Somat. Cell Mol. Genet. **23**:233-236.

FIGURE LEGENDS

Figure 1. Schematic alignment of cDNA sequences isolated by the cDNA library screening and 5' RACE with human *ING1* cDNA. **A**. Mouse *ing1* cDNA, isolated from the library of senescent mouse embryonic fibroblasts, aligned with human *ING1* cDNA. Open reading frame for the human *ING1* clone is indicated. **B**. Alignment of mouse and human *ING1* cDNAs with the products synthesized by 5'-RACE. The position of the gene-specific PCR primers used for the RACE reaction is indicated in panel A.

Figure 2. Analysis of *ing1* transcription by northern hybridization. *Ing1* expression in mouse liver heart and testis was analyzed using a Multiple Tissue Northern blot (CLONTECH laboratories) which was hybridized with the probe corresponding to the common part of *ing1* (1), the 5' end of isoform 1a (2), the 5' end of isoform 1b (3) or the 5' end of isoform 1c (4). Probes were obtained by PCR using *ing1*-specific primers. Arrow shows the position of the RNA marker, 2.37kb in size.

Figure 3. Mapping of the *ing1* exons in the mouse genomic region containing *ing1* cDNA-related sequences. Three *ing1*-positive genomic clones were isolated from the mouse genomic library using *ing1* cDNA as a probe. To determine the position of *ing1* exons, DNA was cut with the indicated combination of restriction enzymes and analyzed by Southern blot hybridization with the indicated probes representing the alternative 5' ends of three *ing1* transcripts (1a, 1b and 1c) or the common part of the gene (3'). The results of hybridization in combination with the PCR data and with partial sequencing of the genomic clones allowed us to determine positions of the *ing1* exons.

Figure 4. Determination of transcription initiation and termination sites in the ing1 gene. The upper panel shows genomic map with the positions of the *ing1* exons, and the structure of *ing1* cDNA clones as determined after cDNA library screening and 5'-RACE analysis. The structure of two known human *ING1* cDNA variants is indicated and their relation to mouse *ing1* sequences. The lower panel demonstrates the the final structure of the ing1 gene determined as a result of precise mapping of transcription initiation and termination sites. In the upper panel, the position of the G/C rich region in the 5' end of clone 1b and the poly(A) region in the 3' part of the *ing1* transcripts are indicated, which interfered with the polymerase reaction during the original 5'- RACE.

Figure 5. Sequences of the putative promoter areas of the *ing1* gene. A. Sequence of the genomic region with the putative promoter areas of isoforms 1a and 1b, the start sites of transcription of these two variants are shown by arrows. The initiator sequence of the 1a isoform, including multiple initiation start sites, is underlined; TATA-like sequence for the 1b isoform is shown in bold. Underlined bold sequence is CAAT box positioned about 100bp upstream from the 1b transcript initiation site. Sp-1 binding sites are displayed in bold italic. B. Sequence of the putative promoter area of the 1c isoform, with the underlined initiator sequence overlapping the two transcription start sites. Multiple sp-1 binding sites are also indicated.

Figure 6. Sequences of *ing1* cDNAs and their predicted protein products. Sequences of the alternative 1a, 1b, and 1c first exons as well as the common *ing1* exon are shown.

The first ATG codon of isoform 1b is underlined, as well as the ATG codon in the common exon that is used as the initiation codon for 1a and 1c translation. Stop codons, indicating the end of translation, are also underlined.

Figure 7. Comparison of mouse and human *Ing1* genes. **A.** Alignment of amino acid sequences of mouse *ing1* 1b isoform and its human orthologue. Numbers of amino acids are indicated; numbers in parenthesis indicate the size of the truncated protein product of 1a and 1c *ing1* isoforms. The underlined methionine in the mouse sequence indicates the beginning of the protein product encoded by 1a and 1c isoforms. Asterisks indicate identical amino acids, while dots indicate conserved changes in amino acid sequence. **B.** Alignment of the 3'-untranslated regions of human and mouse *Ing1* mRNAs. The position of anti-ING1 GSE sequence (5) is underlined.

Figure 8. A multiple alignment of P33ING1, its yeast homologs, and additional PHD finger-containing proteins. The aligned conserved blocks are separated by variable spacers whose length is indicated by numbers; for the N-terminal block, the distance to the protein N-termini is indicated (the sequence of the human paralog of P33ING1 is incomplete). Consensus1 shows amino acid conservation in the P33ING1 family of proteins; consensus2 shows the conservation in an expanded set of PHD finger proteins (in addition to the 6 proteins of the P33ING1 family, the sequences that aligned with P33ING1 with a probability of a random match below 10^{-4} in the first iteration of the PSI-BLAST analysis were included). Each consensus shows amino acid residues conserved in all sequences of the respective set; h indicates a hydrophobic residue, p indicates a

polar residue, "-" indicates a negatively charged residue, s indicates a small residue, and a indicates an aromatic residue. The metal-chelating residues in the PHD finger domain are shown by inverse typing. Each protein is identified by a gene name followed by species abbreviation and the Gene Identification number from the non-redundant protein database at the NCBI. P33ING1p is the human paralog of P33ING1, Mi-2 is a human autoantigen, CHD3 is a chromatin-associated helicase of the SNF2 family; the remaining proteins are uncharacterized gene products designated by their systematic gene names. Species abbreviations: Hs, Homo sapiens, Sc, Saccharomyces cerevisiae, Sp, Schizosaccharomyces pombe, Ce, Caenorhabditis elegans.

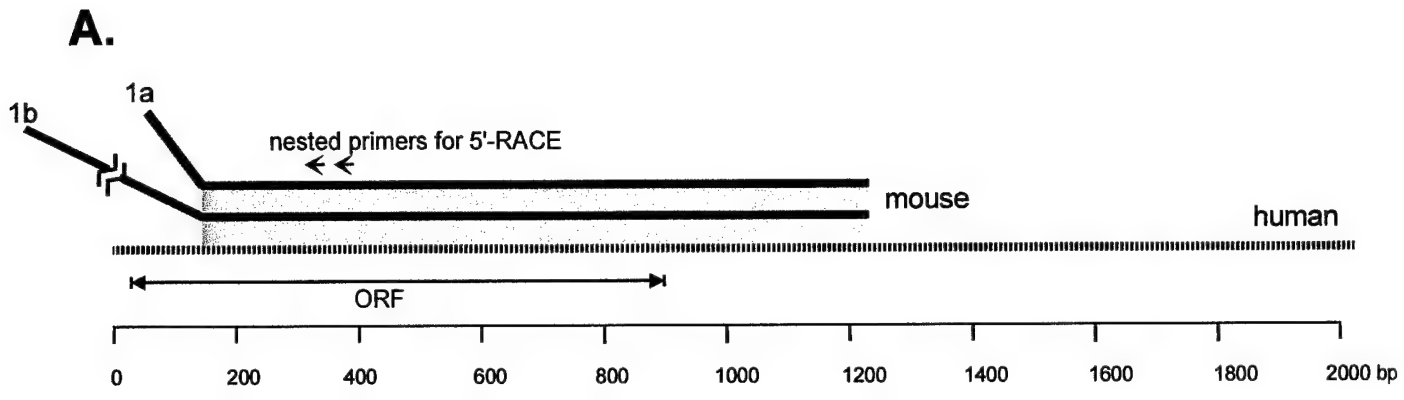
Figure 9. *ing1* expression in adult mouse tissues and embryos analyzed by northern hybridizations. A. Hybridization result with the indicated probes with total mouse RNAs isolated from the indicated organs; photograph of ethidium bromide stained gel is shown as loading control. B. Northern hybridization with the same probes of RNA on CLONTECH's mouse embryonic Multiple Tissue Northern blot. Here, mRNA was isolated from mouse embryos at 7, 11, 15, and 17 days of development. Arrows indicate positions of the 1.35kb, 2.37kb, and 4.4kb RNA markers as well as the positions of 28S and 18S rRNAs.

Figure 10. *ing1* expression in mouse embryos analyzed by *in situ* histo-blot hybridization. Histo-blots were prepared from whole body sections of 10-, 12-, 16- and 18-day embryos, and hybridized with the sense and antisense *ing1*-specific ^{33}P-labeled

RNA probes corresponding to the PHD domain. Antisense actin probe was used as an internal standard.

Figure 11. *ing1* mRNA expression varies depending on cell growth conditions. RNA was isolated from dividing 10(1) (a) and NMuMG (e) cell lines, as well as from the dividing, young mouse embryonic fibroblasts (h). RNA was also isolated from contact-inhibited 10(1) (b) and NMuMG cells (f); serum starved 10(1) cells (c) (64 hours at 0.5% FBS); γ-irradiated 10(1) cells (d) and NMuMG cells (g); and senescent MEFs (i). Filters were probed with total *ing1* probe. Bottom panels show RNA gels from which northern blots were made. Arrows indicate positions of the 2.37kb RNA marker (upper panels) or 28S and 18S rRNAs.

Figure 12. Expression of *ing1* protein products at depends on cell growth conditions. Cell lysates analyzed by western with antibodies against $p33^{ING1}$ were isolated using RIPA buffer from the cells treated in the same way as explained in Figure 11.



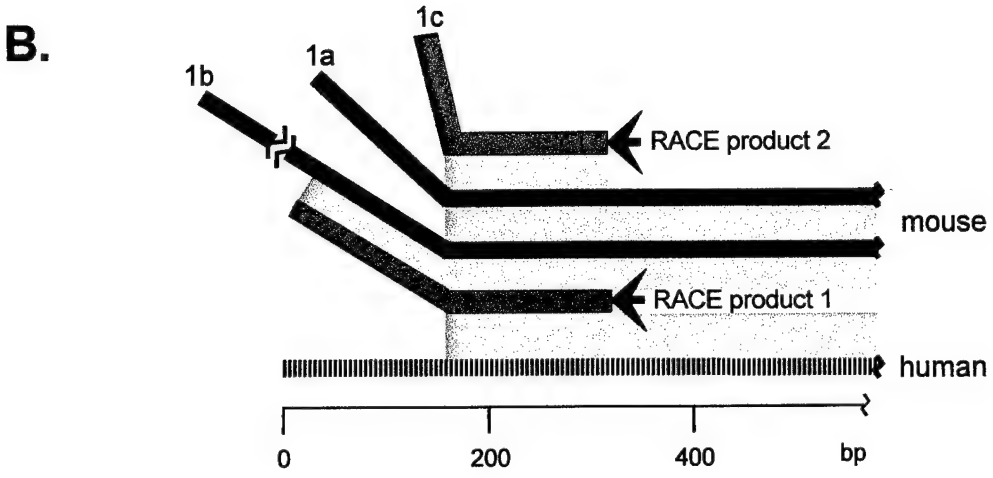


Figure 1 (Zeremski et al.)

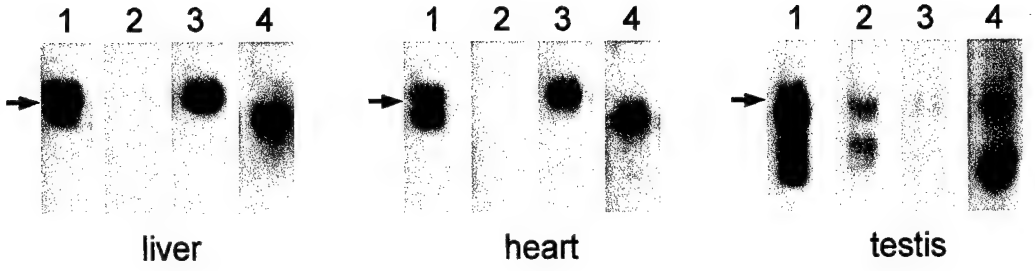


Figure 2 (Zeremski et al.)

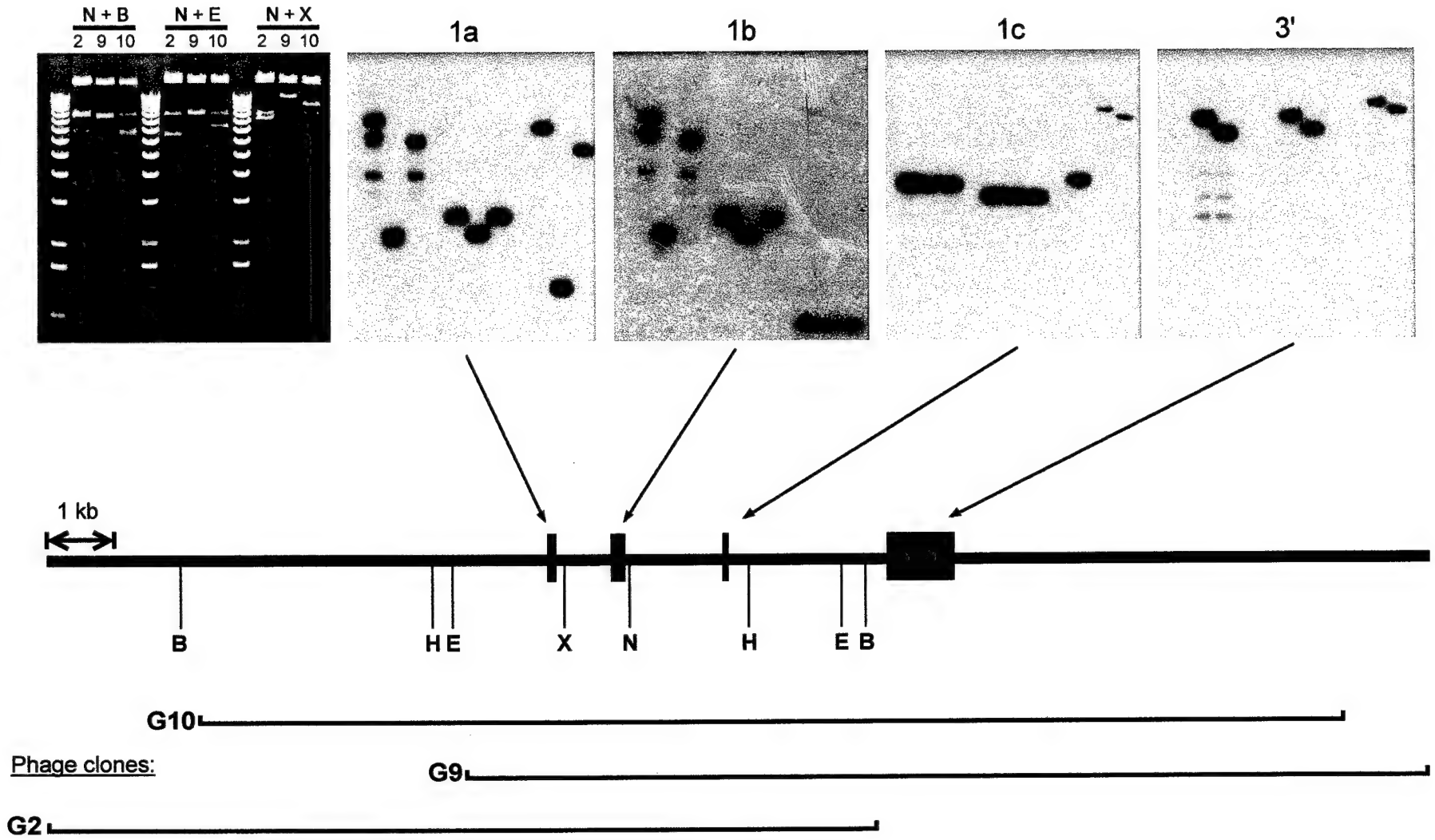


Figure 3 (Zeremski et al.)

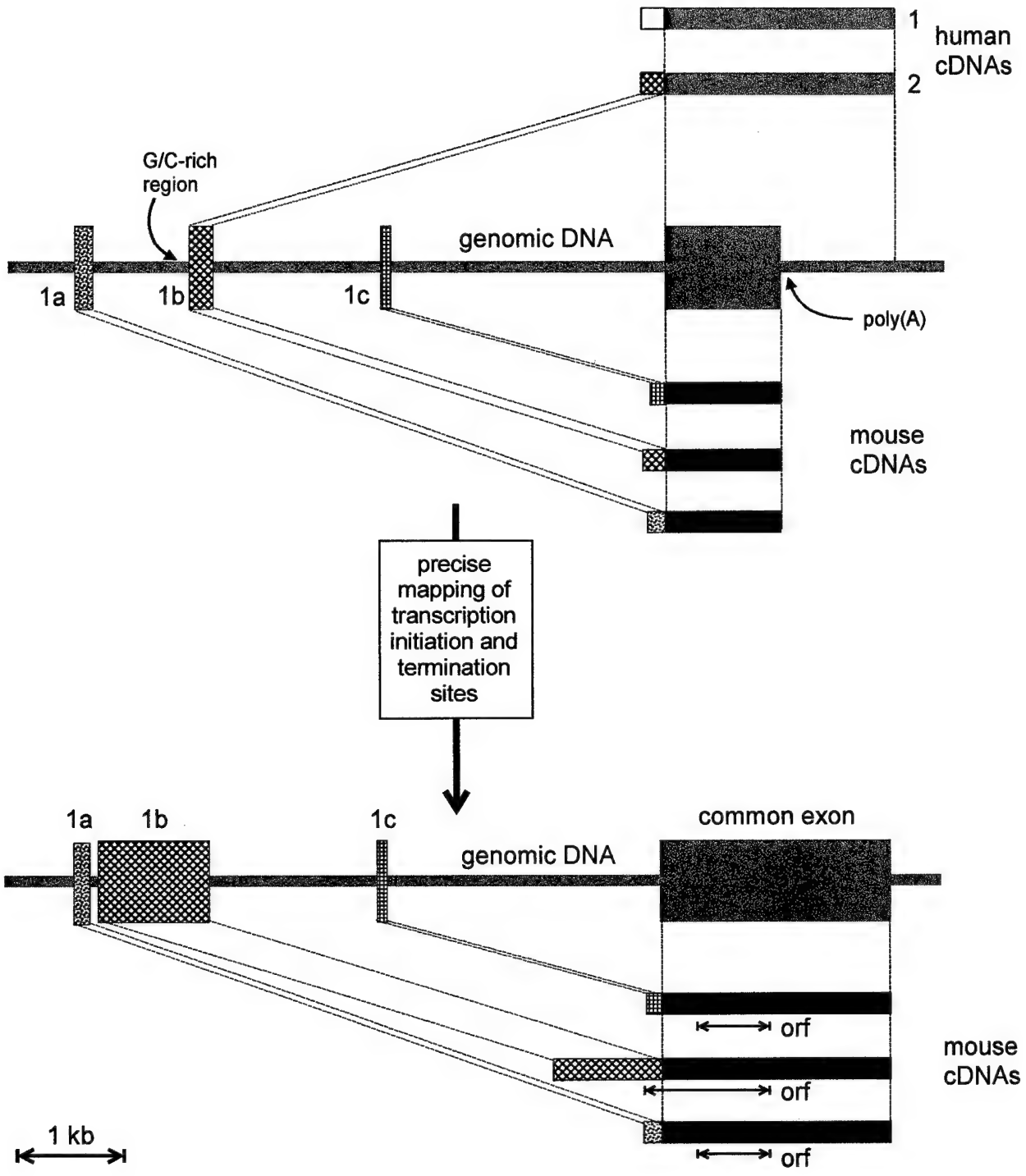


Figure 4 (Zeremski et al.)

A

CACAGTCCCA GCCATACAAA GGCGCGAATG TGGCGACCAC AGGCCAGGCT ***CCGCCC***CA***GG***
GCGGTGCCTA GGCAGGAAGA ***GGGCGG***GACC GCCGGGGCCC GTCGCCTAGG CAACCCCTCG
CTCTGCCCGG GCTTTTCAGA AAGCCCTAAA TCCGGCAGCT GTGGGGGAA***G*** ***GGCGG***GATGG
GAAGCTGAAG GTTGTGCTCC CCTGCGGGAA GCACACTGCG GTGCCCGCTG CCAACTCAGT
GTGACGAGGA CATGGGCAGC CTCAGATGCG GCAGGGAGGG GAAGAGCTAG TTAGGATCAG
AGAAGAGCTG CTCCCGCCAG C***CCGCCC***ACC GTAGATCTCC GCGCTGAAAA GAGCACCCCT
CGGGGGCCCC TTTGTCTCCA GGCCATTCTA AACCCAGCAC CGGGAGGCGA CACAAAGGGA
CGATGG***GGGC GG***AAGGCGCA GGCGCGTGGA GGCGCCGAGG ATGCTGGGAG TGGTGGTCCC
TTTGCGTGTC GATAGCTGCC CTCCCGCGCG GGTCCGCATC CATCGCCAGG GAGCTGCGCC

1a

TGCGCGTCC**A** **CTT**CCG**C**CTC GCTCGGCTCC GCCCGCTCGT CCCTCCTGGA AGAGGACAGC
TCTCTAGTAC TAGGCGTGCG GGGGGCGTTG CCGATCCCAC CACACCTCCT TCTCGTCC***AG***
ATTGGCCGCA GAG*GT*TAGAG TCGCAAGGGG ATTGGCTGAC GCCGTTGCCA TTTCTGTTCG

exon ← ***intron***

1b

GGGGCCC**TAT** **TTATA**CCGCG CTAGGAGGCG GGGGAATGCA GTTAACCCAG CGGAGGGTGG
CTCTTCTCAG GTTAGTCGGT TTGAAGGCAG CGATCTGGAG GCTCGTCCGG CCCCGGCGTG
TCTCGAGGAG ACAGGCTAGC ACCCCGCGGT TCCCGGCCGG CGTGCGGCTC CGCCCAGTAG
GTCCGCCACC GAGTGTTAAA CTCCTAGTAA AGTTTCGCGT CGCCGGTCTC CCCTGCCCAG
GCGACCCCGC CGCTGTCCCC TCGGCGACCC TGGCCCCCAG CGACAATGAC AGGCCACGCC
CCCGCGCGGG CCGGGCTCGC GGCCCGCCGC CCCCCGGCCC GGGACGGTGA GGGGCGTGAA
TGCGGTGGGG GCGGGGCCGC CTCCGGGAGG AGGGGTGGCG GAGCGCATGC GCGCTGCGCG
CGGGGCTGAA TGTTTCCCAA GTGTTTGAAA CTGGTATTTG GGTTTTCCAC GTTGGACAAG
TGCGGCGAGG C*GGGCGGC*GG AGCGCGCCCT TCCCGCAGCC GGCCCCGCTC TCTCCGCTCT
CCACGCTGGT GGCGGTGTGG GCGGCGGTGG CCGGCCCTCT TTGGGTGTGT GCGCCCTAGT
ACGCACGGCC CCCCGCGCCG CCCGCCGCCC GGGAGGGGGC CTGCACGGCC GGCGGGGTGT
GCGCTTGAGC TTCGGCCGCC GCGGCCCGCG CCTGCAGGCG CCGGGGTCCC CGCGGCCCCG
GAGGCGGCGC GGCAGGCGCC AGATGTAGCC GCCGGCCAGC CCGGCCGGAG CGGCGGGGGG
GCGCGGAAGG GCGAGAGAGC TTTGCATTTT GCAGTGCTGT TTGAGGGGGG CGGGGGGTGG
AGGAAGCGGA AAGCCGCCGA ATCGCCGGGG ACCTCCGGGG TGAACCATGT TGAGTCCTGC
CAACGGGGAG CAGATCCACC TGGTGAACTA TGTGGAGGAT TACCTGGACT CAATCGAGTC
ACTGCCTTTC GACCTGCAGA GGAACGTCTC GCTGATGCGG GAGATCGACG CCAAATACCA
AGGTACGGCG

exon ← ***intron***

B

ACAGAGTCAT TGGG***GGGCGG*** GCAGTATGG***G*** ***GGCGG***GCTTT GCTCGGGACT CGCTCCAGTC
TGTGTTTCCG CTGTCTGCTT TTTTTTTTTT TTTTTTCCTC TGGTAGAGGG CC***GGGCGG***AG
AGAGGAGGTG AGTTGATTTG AATGTCTTCG GGTCGCCCGT CCTCTGGCCT TGGGTTGGCT
CCTGTCGCTG CTG***GGGCGG***G CCACCGTCGC TGGCTCTGCG CGCTGATTGG TGTACCTCCT
GGTCT***CCGCC*** ***C***TCAGCGTCG TCGACTCACA TAGGCTTC***GG*** ***GCGGGGCGG***G GCA***GGGCGG***G

1c

GCGCGCTTCC TGAGTCTCGT AGGCTGGAGT GGATCGCGGC **CA**CTTTCCGG CTGCGAGGCT
ATGGCGGCGG TGGCCTCCGG GAGGATGCTG CGCTTACCTT CTCTGCTCTG GCTCCCCGCG
GGAGCCTCTG ATCGCTTGTC GCGTTTCCGG TAGGCTTGAA TGAGCGGGAG TGCCTCTGTC

exon ← ***intron***

Figure 5 (Zeremski et al.)

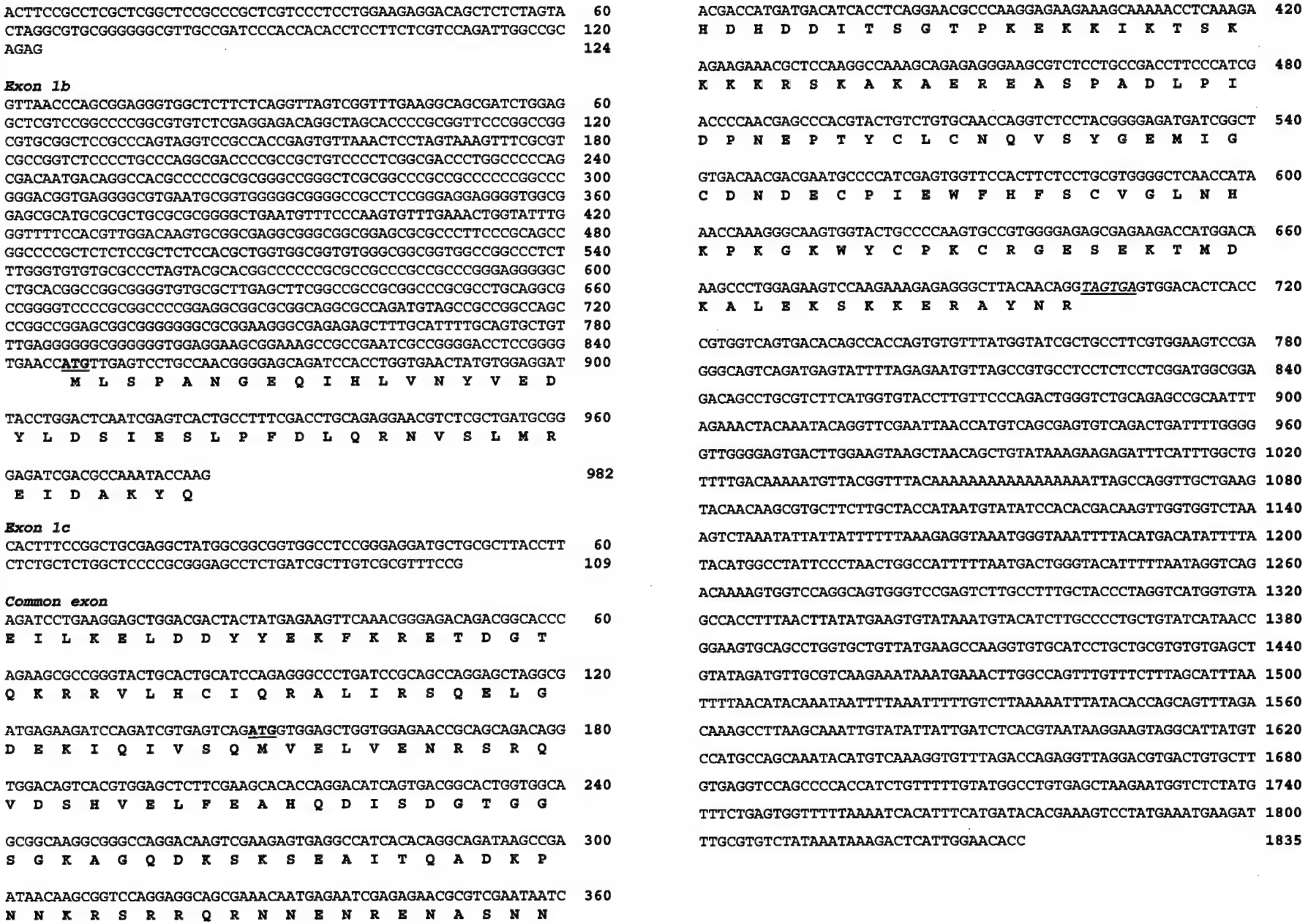


Figure 6 (Zeremski et al.)

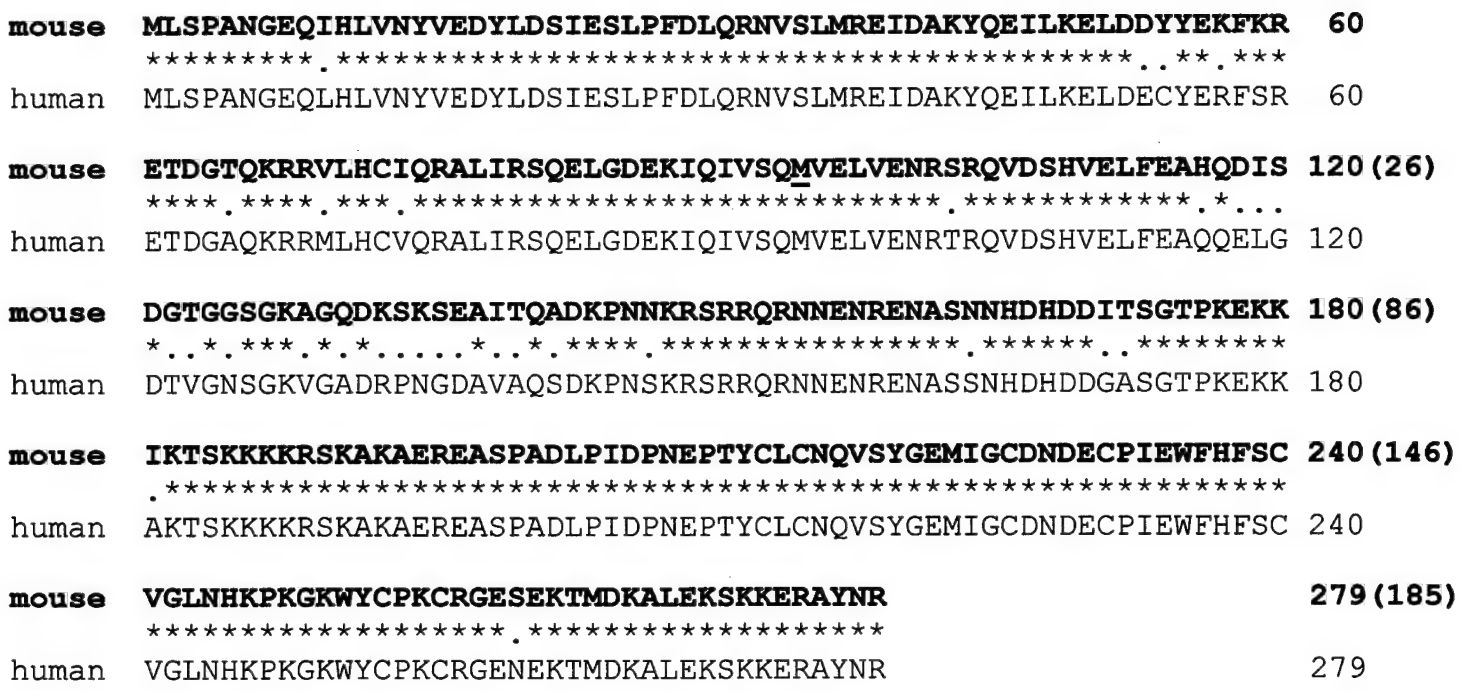


Figure 7A (Zeremski et al.)

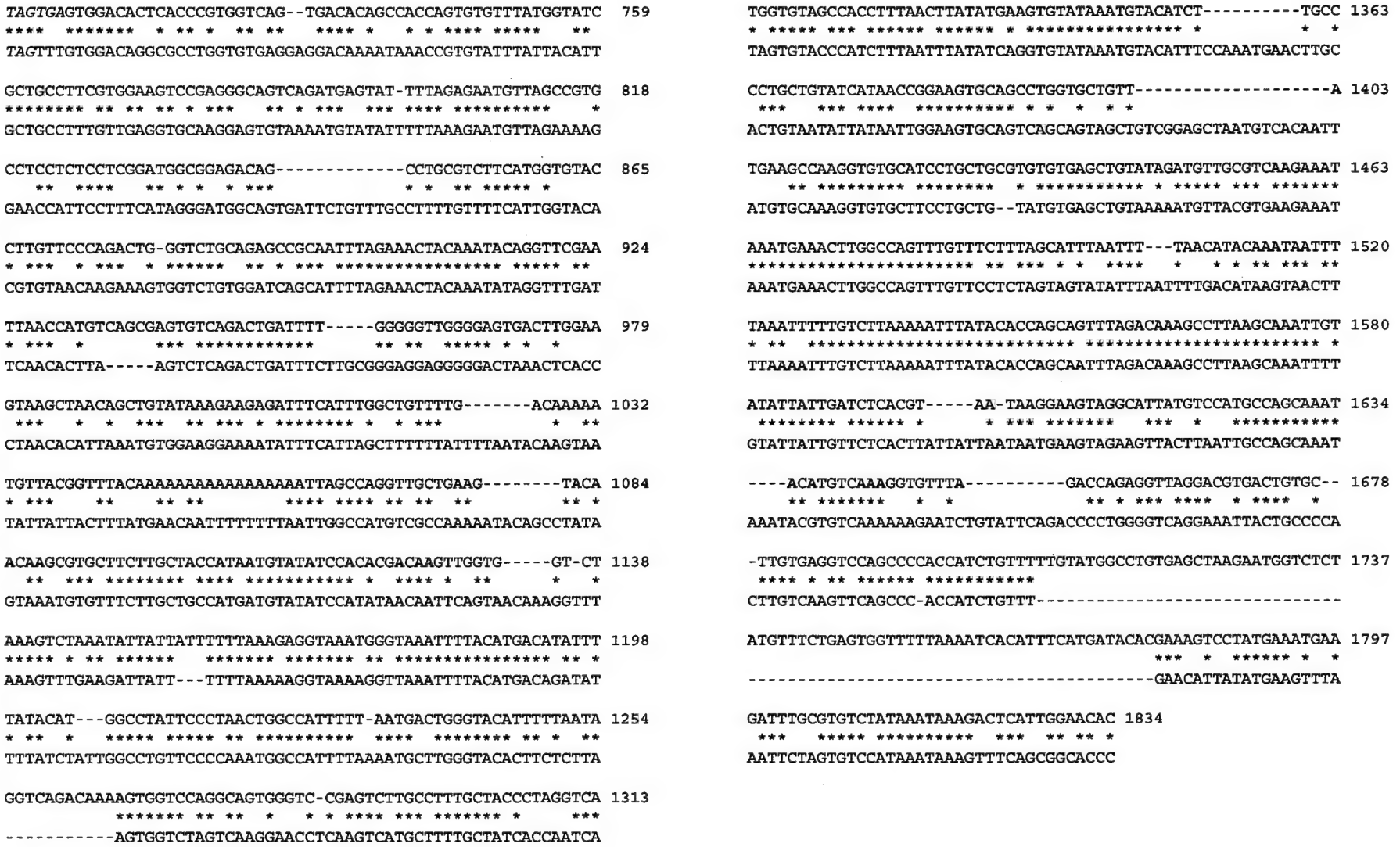


Figure 7B (Zeremski et al.)

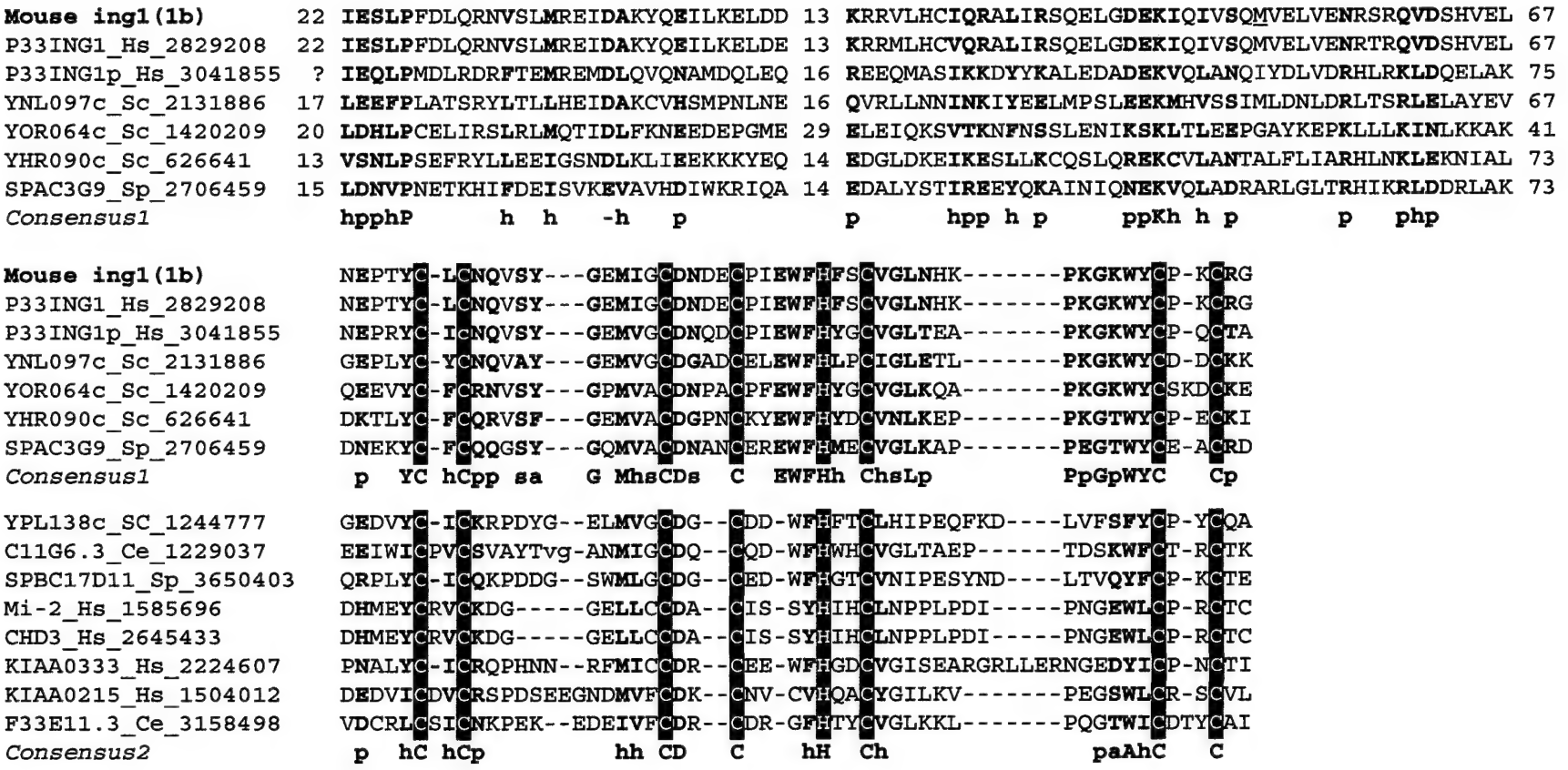
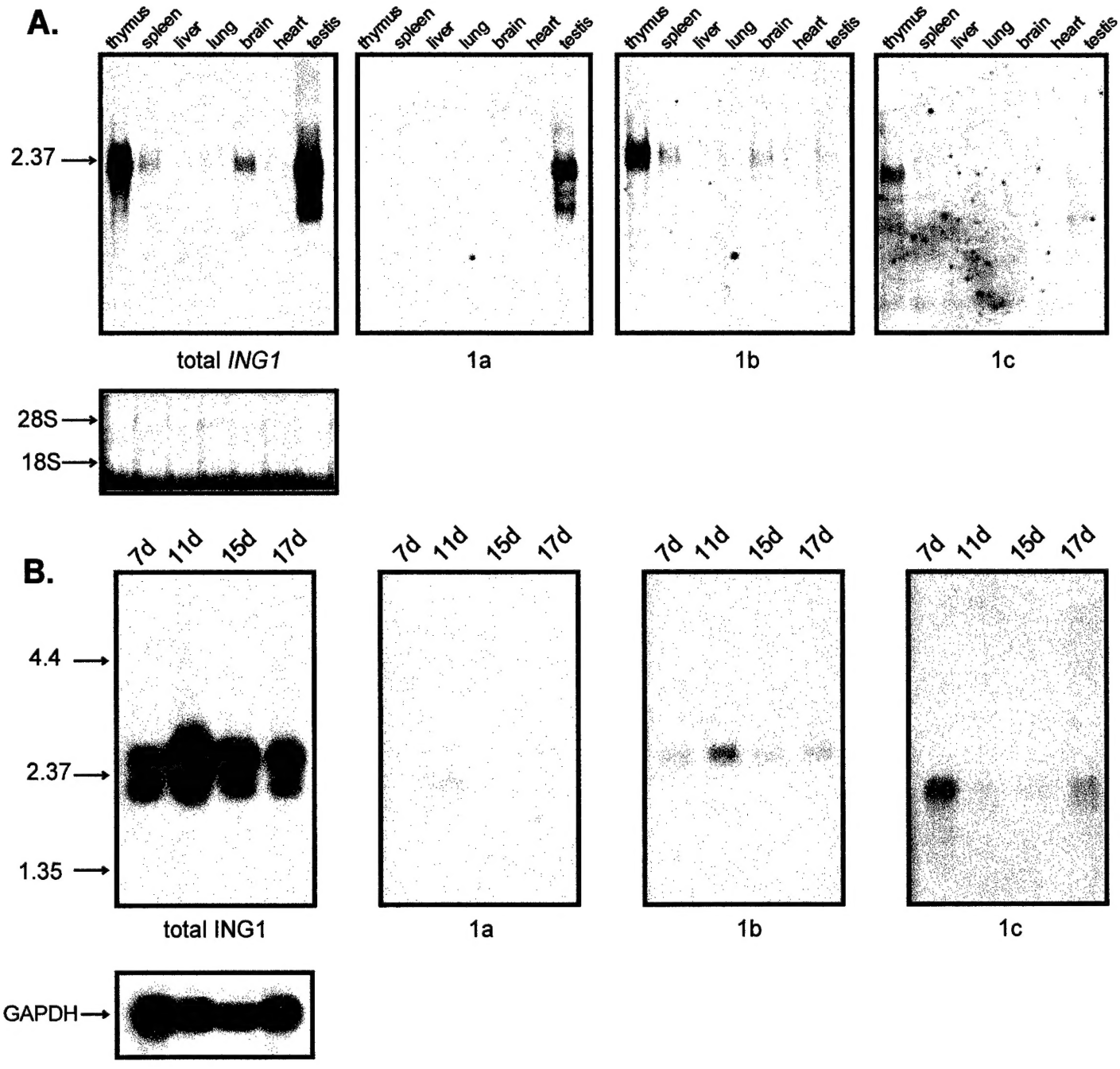


Figure 8 (Zeremski et al.)



Fifure 9 (Zeremski et al.)

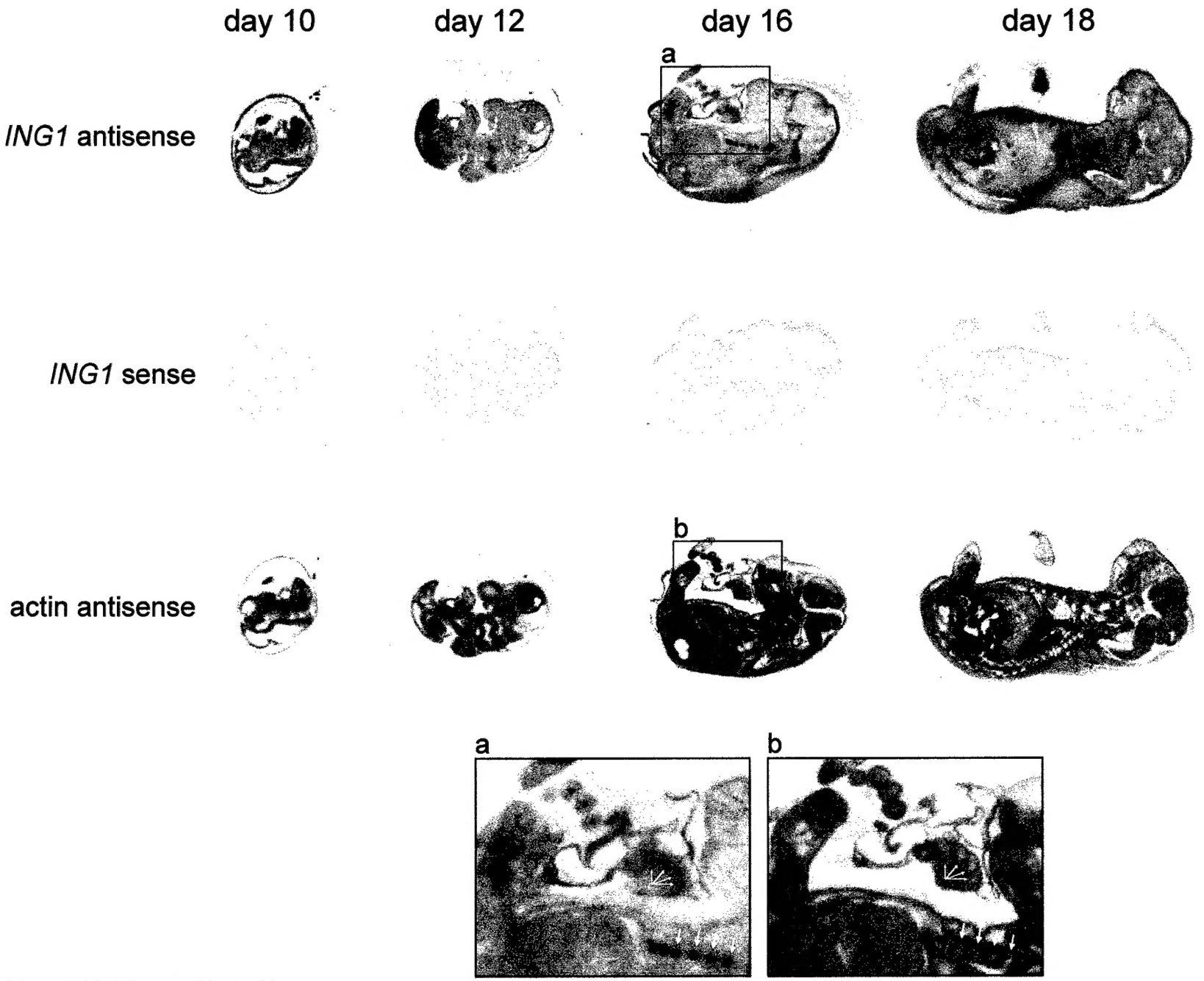


Figure 10 (Zeremski et al.)

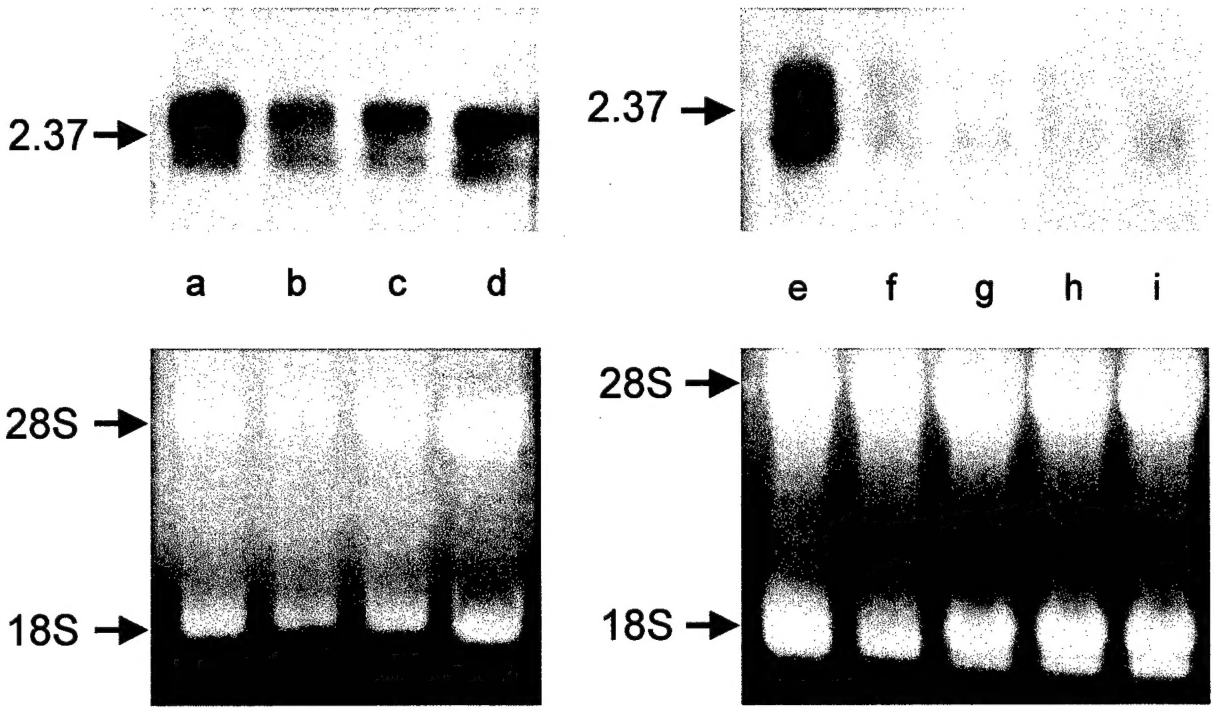


Figure 11 (Zeremski et al.)

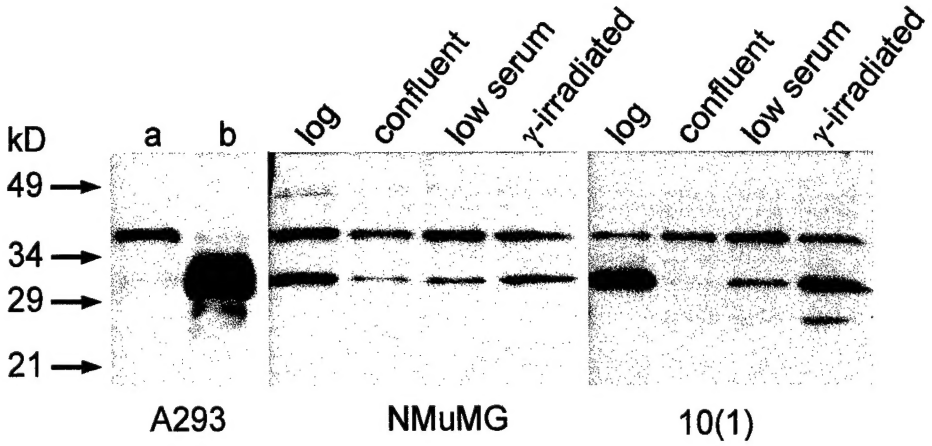


Figure 12 (Zeremski et al.)